Multi-Level Repetition Benchmarking

Marek Cuchý

marek.cuchy@agents.fel.cvut.cz

B4M36ESW

March 4, 2019

AI CENTER

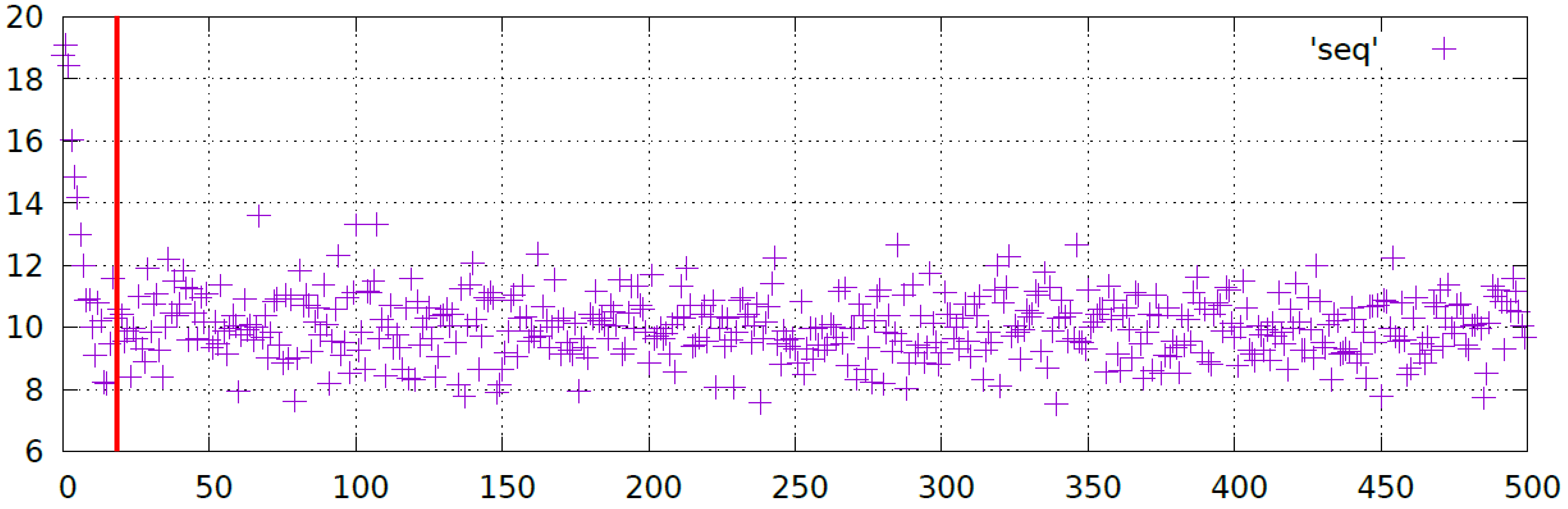
CTU
CZECH TECHNICAL UNIVERSITY IN PRAGUE

Multi-Level Repetition

- Variance in measurements may occur at higher levels
  - ➔we need to repeat measurements at least on the level of the variance
- The highest level is the most important one
- Levels:
  1. Iteration – loop body
  2. Execution – running of the program
  3. Compilation (stable in Java)

Warm-up

- Measurements usefull only after reaching a steady/independent state
- E.g. by manual inspection of sequence plot



Execution Time with Confidence Interval

$$\overline{Y} \quad \pm \quad t_{1-\frac{\alpha}{2},\nu}\sqrt{\frac{1}{r_n(r_n-1)}\sum_{j_n=1}^{r_n}\left(\overline{Y}_{j_n\underbrace{\bullet\cdots\bullet}_{n-1}}-\overline{Y}\right)^2}$$

Execution Time with Confidence Interval

$$\overline{Y} \quad \pm \quad t_{1-\frac{\alpha}{2},\nu}\sqrt{\frac{1}{r_n(r_n-1)}\sum_{j_n=1}^{r_n}\left(\overline{Y}_{j_n\underbrace{\bullet\cdots\bullet}_{n-1}}-\overline{Y}\right)^2}$$

- $(1-\alpha)$ – confidence interval (e.g. 95% confidence → $\alpha = 0.05$)
- n - number of levels
- r_n - number of repetition on the highest level
- $\overline{Y}$ - mean across all measurements (after reaching independent state)
- $t_{1-\frac{\alpha}{2},v}$ - $\left(1-\frac{\alpha}{2}\right)$-quantile of the t-distribution with $v = r_n - 1$ degrees of freedom, can be found in a table

Execution Time with Confidence Interval

$$\overline{Y} \quad \pm \quad t_{1-\frac{\alpha}{2},\nu}\sqrt{\frac{1}{r_n(r_n-1)}\sum_{j_n=1}^{r_n}\left(\overline{Y}_{j_n\underbrace{\bullet\cdots\bullet}_{n-1}}-\overline{Y}\right)^2}$$

Execution Time with Confidence Interval

$$\overline{Y} \quad \pm \quad t_{1-\frac{\alpha}{2},\nu} \sqrt{\frac{1}{r_n(r_n-1)} \sum_{j_n=1}^{r_n} \left(\overline{Y}_{j_n \underbrace{\bullet \cdots \bullet}_{n-1}} - \overline{Y} \right)^2}$$

- The mean across all measurements but the highest level for the j_n-th repetition

Execution Time with Confidence Interval

$$\overline{Y} \quad \pm \quad t_{1-\frac{\alpha}{2},\nu} \sqrt{\frac{1}{r_n(r_n-1)} \sum_{j_n=1}^{r_n} \left(\overline{Y}_{j_n \underbrace{\bullet \cdots \bullet}_{n-1}} - \overline{Y} \right)^2}$$

- The mean across all measurements but the highest level for j_n-th repetition
- Example: $n = 2, r_1 = 10, r_n = r_2 = 20$

$$\overline{Y}_{5 \underbrace{\bullet \cdots \bullet}_{2-1}} = \frac{\sum_{i=1}^{r_1} Y_{5,i}}{r_1}$$

$Y_{a,b}$ - measured time of the b-th iteration of the a-th execution

Speed-Up Ratios with Confidence Interval

$$\frac{\overline{Y} \cdot \overline{Y'} \mp \sqrt{\left(\overline{Y} \cdot \overline{Y'}\right)^2 - \left(\overline{Y}^2 - h^2\right)\left(\overline{Y'}^2 - h'^2\right)}}{\overline{Y}^2 - h^2}$$

$$h = \sqrt{t^2_{\frac{\alpha}{2},\nu} \frac{S_n^2}{r_n}} \qquad h' = \sqrt{t^2_{\frac{\alpha}{2},\nu} \frac{S'^2_n}{r_n}}$$

- $\overline{Y}, \overline{Y'}$ - means of the compared implementations
- h, h' - half-widths of the confidence intervals for the single implementations
- S_n^2, S'^2_n - variance estimator of the n-th level

Speed-Up Ratios with Confidence Interval

$$\frac{\overline{Y}\cdot\overline{Y'} \mp \sqrt{(\overline{Y}\cdot\overline{Y'})^2 - \left(\overline{Y}^2 - h^2\right)\left(\overline{Y'}^2 - h'^2\right)}}{\overline{Y}^2 - h^2}$$

$$h = \sqrt{t^2_{\frac{\alpha}{2},\nu}\frac{S_n^2}{r_n}} \qquad h' = \sqrt{t^2_{\frac{\alpha}{2},\nu}\frac{S'^2_n}{r_n}}$$

- $\overline{Y}, \overline{Y'}$ - means of the compared implementations
- h, h' - half-widths of the confidence intervals for the single implementations
- S_n^2, S'^2_n - variance estimators of the n-th level ???

Speed-Up Ratios with Confidence Interval

- S_n^2, S'^2_n - variance estimators of the n-th level

  → Don‘t have to be calculated

Speed-Up Ratios with Confidence Interval

- S_n^2, S'^2_n - variance estimators of the n-th level

  → Don‘t have to be calculated

$$t_{1-\frac{\alpha}{2},\nu}\sqrt{\frac{1}{r_n(r_n-1)}\sum_{j_n=1}^{r_n}\left(\overline{Y}_{j_n\underbrace{\bullet\cdots\bullet}_{n-1}}-\overline{Y}\right)^2} \overset{[1]}{=} t_{1-\frac{\alpha}{2},\nu}\sqrt{\frac{S_n^2}{r_n}}$$

equality from [1]

Speed-Up Ratios with Confidence Interval

- S_n^2, S'^2_n - variance estimators of the n-th level

  → Don't have to be calculated

$$t_{1-\frac{\alpha}{2},\nu}\sqrt{\frac{1}{r_n(r_n-1)}\sum_{j_n=1}^{r_n}\left(\overline{Y}_{j_n\underbrace{\bullet\cdots\bullet}_{n-1}}-\overline{Y}\right)^2} \overset{[1]}{=} t_{1-\frac{\alpha}{2},\nu}\sqrt{\frac{S_n^2}{r_n}} = \sqrt{t^2_{\frac{\alpha}{2},\nu}\frac{S_n^2}{r_n}} = h$$

- t-distribution is symmetric → $t^2_{\frac{\alpha}{2},v} = t^2_{1-\frac{\alpha}{2},v}$

Speed-Up Ratios with Confidence Interval

$$h = t_{1-\frac{\alpha}{2},\nu}\sqrt{\frac{1}{r_n(r_n-1)}\sum_{j_n=1}^{r_n}\left(\overline{Y}_{j_n\underbrace{\bullet\cdots\bullet}_{n-1}}-\overline{Y}\right)^2}$$

References

[1] Kalibera, T. and Jones, R. E. (2013) Rigorous Benchmarking in Reasonable Time. In: ACM SIGPLAN International Symposium on Memory Management (ISMM 2013), 20–12 June, 2013, Seattle, Washington, USA. http://kar.kent.ac.uk/33611/